2010-12

A security perspective of transitioning organizations to the DoD cloud

Dinolt, George W.

Monterey, California. Naval Postgraduate School

NPS-CS-10-012

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

A Security Perspective of Transitioning Organizations to the DoD Cloud

by

A. J. Nelson, J. B. Michael, G. W. Dinolt

December 2010

Approved for public release; distribution is unlimited

Prepared for: Office of the DoD Chief Information Officer
1851 S. Bell St., Suite 600
Arlington, VA 22202

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

Daniel T. Oliver
President

Leonard A. Ferrari
Executive Vice President and
Provost

This report was prepared for and funded by the Office of the DOD CIO.

This report was prepared by:

Alex Nelson
Research Assistant
Naval Postgraduate School

James Bret Michael
Professor of Computer Science and
Electrical Engineering
Naval Postgraduate School

George W. Dinolt
Professor of Practice in Cyber Operations
Naval Postgraduate School

Reviewed by:

Peter J. Denning, Chairman
Department of Computer Science

Released by:

Douglas J. Fouts
Acting Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
10 Dec 2010	Technical Report	Jan 18 – Dec 10, 2010

4. TITLE AND SUBTITLE
A Security Perspective of Transitioning Organizations to the DoD Cloud

5a. CONTRACT NUMBER
DWAM00390

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Alex J. Nelson, James Bret Michael, George W. Dinolt

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Naval Postgraduate School
1411 Cunningham Road,
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-CS-10-012

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of DoD Chief Information Officer
1851 S. Bell Street
Suite 600
Arlington, VA 22202

10. SPONSOR/MONITOR'S ACRONYM(S) DoD CIO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
The views expressed in this report are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

14. ABSTRACT

Security consistently emerges as a top concern with IT administrators and users when surveyed on transitioning to a cloud infrastructure. To initiate cloud usage, security concerns must be addressed—however, security and improved usability will be what keeps the cloud around. We address in this paper security and usability issues and opportunities with the Department of Defense transitioning to a cloud infrastructure.

The major contributions of this paper are three enablers for a secure and usable DoD Cloud infrastructure. We also contribute a transition plan for organizations in terms of four milestones in cloud adoption, assuming that the necessary secure and usable Cloud infrastructure exists at each milestone. We present three perspectives throughout the paper: The DoD in designing its Cloud, industry efforts to support the Cloud, and organizations joining the Cloud.

15. SUBJECT TERMS
Cloud computing, object and content management, document sharing, cloud infrastructure, security, usability

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					James Bret Michael
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
Unclass-ified	Unclass-ified	Unclass-ified	UL	28	831-656-2655

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

Abstract

Security consistently emerges as a top concern with IT administrators and users when surveyed on transitioning to a cloud infrastructure. To initiate cloud usage, security concerns must be addressed—however, security and improved usability will be what keeps the cloud around. We address in this paper security and usability issues and opportunities with the Department of Defense transitioning to a cloud infrastructure.

The major contributions of this paper are three enablers for a secure and usable DoD Cloud infrastructure. We also contribute a transition plan for organizations in terms of four milestones in cloud adoption, assuming that the necessary secure and usable Cloud infrastructure exists at each milestone. We present three perspectives throughout the paper: The DoD in designing its Cloud, industry efforts to support the Cloud, and organizations joining the Cloud.

THIS PAGE INTENTIONALLY LEFT BLANK

1 Overview

The DoD is investigating the issues involved in migrating much of the information processing on the NIPRNET to a cloud or multiple cloud infrastructure. The assumption is that the data and applications for the approximately 3.5 million users would have to be migrated to this new infrastructure. The assumption is also that the management and infrastructure savings would more than offset the costs of such a move, while providing security at least equivalent to the current NIPRNET. In this paper we investigate some of the security issues that would arise and offer some suggestions as to how to address them.

The DoD Cloud (referred to as "The Cloud") effort will provide a common storage and application space to organizations within the DoD. Today, this worries a lot of users and IT administrators. Only a minority is comfortable with entrusting workflows to a cloud service parented by any organization, and hosting data in an offsite store raises many concerns. Security is consistently the top concern–putting all the documents within a massive, off-site store provokes fears of leaks, unavailability, or tedium in accessing data and services.

The Cloud must supply fundamental ("Necessary, though not necessarily sufficient") security policy. When organizations employ the Cloud, the data they place on it and the practices of use must adhere to its security policies. Getting all organizations' policies to work together otherwise is a highly difficult and politics-laden task that grows in personnel (administrative negotatiors') work time quadratic to the number of organizations joining[1]. All this work would come from getting organizations to agree on basic security policies, some of which may contradict each other. For instance, user and data management have many universal security components, users having roles and files having permissions. Placing the common denominator of security policies within the Cloud saves on negotiations and allows for basic Cloud services. For instance, all user and data management should happen through the Cloud.

Good security policy can help prevent surreptitious access to the central document repository, and be an integral part of preventing damaging leaks such as the Wikileaks incidents [16]. Defining that policy and its resulting interface is one of the central challenges in Cloud deployment and adoption. As security should be integral to a system design, especially at the beginning, we describe here development of the Cloud with an emphasis on security aspects. As usability is integral to good security, we also describe how the Cloud can enable usable security.

We describe a particular milestone state as a goal. This state may not be the last state, or ultimate vision of an organization operating "On the Cloud." However, it is significant in its departure from the present state of infrastructure management, and in its attempt to ease security interactions.

A major goal of the Cloud is to reduce the *Email problem*: email is the most convenient

[1]The graph-theoretic representation is as follows: Consider a graph where the nodes are security policies and the edges between nodes are successful compositions between security policies. Completing this graph means composing $\frac{n(n-1)}{2}$ policies. n will likely be one or more per organization joining.

file-sharing mechanism today. Email has significant downsides traded for ease of use: It removes all technical restrictions on further file propagation; and it generates redundant, immutable copies of files, wasting space. We intend for the Cloud to reduce the incentives to use this transfer channel.

From the security infrastructure perspective, this milestone will require significant advances from the present state of the art in several areas integral to security and everyday work: authentication, storage, and applications. Designing the Cloud for 3.5 million users requires distributed data centers and user administration, so we will discuss some of the technical and policy security implications of this distribution.

1.1 Authentication

A fundamental concern in security is deciding who is trusted to do what. This question is unanswerable without a sound method of presenting an identity. Organizations have their own methods of proving identities and capabilities, and their own security labels, but outside organizational boundaries those policies and labels have to be translated. This problem, the *domain composition* problem, has significant administrative and technical performance issues discussed in Section 2.3.

1.2 Storage

The Cloud should be geographically distributed. There are several reasons for this, most prominent among them network bandwidth problems and storage capacity. We do not address networking issues here, but one simple perspective is thinking what happens when the West Coast arrives at the office and has to pull their virtual desktops in from the Carolinas – there is a better way.

The storage issue is one of sheer mass of users. The storage must be split because general storage at magnitudes of over 10^4 users is presently unstudied. Leung *et al.* profiled storage for a company that had about 1,000 users using a general-purpose storage service (users were within the "corporate" groups of the company, versus the "engineering" groups) [12]. Storage at the scale of 10^7 users presents such significant issues in data and metadata management that it will be more feasible to simply distribute the Cloud's storage to take advantage of naturally existing delineations, such as organizational boundaries and geography.

1.3 Applications

Cloud applications present a significant security challenge, in no small part due to their sources. While there will be some applications that act like web applications, some software executing on the Cloud will be user-supplied. Those user-supplied applications will be necessary to perform arbitrary computations on unstructured data, and there is the risk: If malicious or faulty software breaks out of virtualization layers, the span of that software

is arbitrary. We must also address what distinct security advantages the Cloud will provide, in the face of the user's ability to supply software and data.

The Cloud offers strong advantages in data integrity over today's DoD computing environment. Centralizing the data stores for some large-scale Cloud applications means there will be structural requirements of the data, so there is already a basic understanding of data integrity. This can be the basis for a Clark-Wilson security model [2], which describes maintaining integrity in the face of application faults and user fraud. Drusinsky *et al.* [3] propose an approach to structuring Cloud data that can greatly aid implementing a Clark-Wilson model and gaining those integrity advantages.

Containing software and data within the Cloud unifies the security framework, and make possible some access control and auditing implementations. One implementation, the managed chain of delegations from user to application to server to data store, will be possible because of occurring within the Cloud. Also thanks to this Cloud envelopment, the opportunity to audit accesses and application flow improves over today. Ultimately, while the risk of consolidating data and applications is high, there are significant improvements to global system security made possible by consolidating the management and policies.

1.4 Perspectives

The rest of this document describes three perspectives of developing and deploying the Cloud. First, we outline the DoD needs and policies to specify a successful migration to the Cloud. Second, we outline what industry efforts must support this migration. Last, we outline how individual organizations transition to the Cloud close to a prior outline of transitioning to the vision of Cloud Computing [4].

2 Security specifications of the DoD Cloud

In our deployment milestone, the Cloud must provide an environment matching NIPRNET's isolation. It provides at least Storage and Software as a Service, through the following interfaces:

- The storage interface appears to the user as a file system, allowing general storage accessible to all users along with specially structured storage accessible by Cloud applications.
- The file system provides a Uniform Resource Identifier (URI) scheme (*e.g.* `cloud://cio-nii.defense.gov/Users/`).
- The file system allows for a mechanism to reference versions of documents, maintaining an inheritable security definition.

- The application interface appears similar to a web portal to applications. The file system interface may not be as appropriate a display mechanism for the portal, as users see different subsets of these applications.

Security is necessary for consolidating DoD data, so the Cloud must provide a security policy to which all organizations can adhere. To date, the most global security policy within the DoD is essentially Discretionary Access Control (DAC) [11]. Unfortunately, revising this policy, for example in implementing Mandatory Access Control, would likely delay deployment. Thus, the initial Cloud deployment must provide mechanisms to enforce DAC, including the following:

- Storage in the form of a large, versioned, distributed file system.
- A single sign-on infrastructure.
- Public Key Infrastructure (PKI) that encompasses at least Cloud member organizations.
- A mechanism to enable simple data sharing.

Data and users associated with an organization are folded in under a *domain*, with the term used as presently understood in representing user identities and managing storage permissions. Miltchev *et al.* surveyed several production and many research file systems [13], focusing on administrative overhead involved for sharing data. Their findings indicate that a 3.5 million user file system with autonomous sharing is beyond the state of the art in production and research file systems.

We propose here an outline of steps to take to resolve the *autonomous delegation problem*, which is a user (or program) delegating their own permission on a resource (*e.g.* reading a file) to another user (or program), without having to involve a security administrator and without violating security policy. Without autonomous delegation, the Cloud's lack of usability hinders adoption drastically, and the most usable file sharing is emailing files as it is today. Email bypasses security definitions on files, and only shares one particular version of a file at a time, so email is a security fault and a trade of storage space (for each version) for collaboration. The Cloud must provide some sharing mechanism to match email in ease, and its architecture can support such a mechanism.

2.1 File system interface

With the Cloud acting as a large network and storage device, the security interface to data can be viewed as the security portion of a file system interface, where the file system is mounted as a network share appropriate to the user's operating system. The NIPRNET isolation will stem from the file system being accessible (or "attachable," or "mountable") only within the Cloud environment. Names for resources within the Cloud should be defined

with URIs, which can include version names. It will be an important part of maintaining data integrity that the Cloud provide data versioning. Security rules, visible as metadata as they are presented in current file systems, will be set through the normal mechanisms such as file system GUI dialogs along with others, one of which we describe in this section.

A security (not usability) ideal would be that data that goes into the Cloud should remain in the Cloud. However, besides working offline, getting data in and out will be necessary at times. Thus, a reasonable mechanism must exist for data ingest and export, and DoD must specify policy on what controls data in- and out-flow. Industry must present the DoD with possibilities for disconnected network operations and trustworthy ingest and export.

2.2 Data sharing mechanism

Johnson *et al.* made an economic analogy of file system security policies [9]. They found that "The failure modes of file systems that enforce centrally-imposed access control policies are similar to the failure modes of centrally planned economies: individuals either learn to circumvent these restrictions as matters of necessity or desert the system entirely, subverting the goals behind the central policy." They noted the most common mechanism to circumvent frustrating security policies is email.

The Cloud should make data sharing at least as easy as emailing files around, as one of its fundamental goals. Since the Cloud is the data repository, and hosts the access control mechanisms and user identity and roles, it should have the freedom of direction to allow users to share information better than prior efforts—and there have been many prior efforts, in production and research file systems. Past file systems have primarily had problems with sharing data across domains, as noted by Miltchev *et al.* [13]. The next section addresses the cross-domain issues.

The mechanism to share files has as its only requirements that it not violate security policies and that it not be so frustrating it makes users resort to emailing files. We propose a mechanism here that fits these requirements, and which the Cloud is well-suited to facilitate.

2.2.1 Proposal for reference attachments

Johnson *et al.* observed that emailing files to share data satisfies many common security expectations of a user [9]. We propose *reference attachments* to embrace this practice with a subtle alteration: Links are sent instead of entire files. This mechanism for attaching references acts as a plug-in to the Cloud email clients. It must recognize Cloud URIs, or generate one from a file reference, and interact with the file system to verify Access Control List (ACL) modification possibilities, adding names and roles to the ACLs. Because of the versioning inherent in the storage, these ACLs can be version-specific or time-independent (*e.g.* setting up work with a collaborator should allow future reads and writes).

Email has been the easiest method of sharing files known to user and frustrated security administrator alike. By sending references to Cloud-stored data instead of files, email is no longer a target for intercepting sensitive files. The reference does not even count as a permission token, vulnerable to repurposing or cumbersome to sign: Access to the file is handled by file system mechanisms and user identity. Intercepting the reference reveals no more than the resource name. The email is merely a signal that the recipient has access permission, and a logistic-handler that makes sure the file is ready for the user's access.

Johnson *et al.* attempted to implement a mechanism [9]. Unfortunately, the authors were unable to implement their mechanism on their target file system for reasons specific to that particular file system and its over-network sharing operations (NTFS). We believe a Cloud file system with more support than a handful of researchers can implement an over-email sharing mechanism, regardless of the file system ultimately used as the basis of Cloud storage.

Industry would have to provide a plugin for each email client to be supported. The important components of this effort would be the usability of the file reference inclusion (most users understand drag and drop pretty well), and the file system security interactions with all of the file access control rules. It must be clear, but unobtrusive, what permissions a user is granting to a file, and to how many versions.

The ability to identify users, perhaps also according to roles, is an important prerequisite to this mechanism. The next section addresses how these users are named.

2.3 Domain composition

Miltchev *et al.* indicate the strongest factor inhibiting autonomous delegation is a lack of trust among domains in representing users and their rights to each other, and the administrative difficulty in establishing that trust [13]. For instance, in a network of domains where all peer domains are equal, trust relations must be established pairwise between each domain to allow interaction beyond unauthenticated email (*e.g.* sharing a file, or running an application). Many of the better-known forms of access control, such as Role Based Access Control (RBAC) and Attribute Based Access Control (ABAC), have significant scalability and consistency issues in cross-domain deployments [10].

Figure 1 shows an example scenario, where a user in a domain D_1 wishes to access a file in domain D_3. If the user were a member of domain D_3, this would be a simple matter of checking file permissions against the user's identifier. However, a user's identifier is only recognized within the domain to which the user authenticates. The user `john.doe@nps.edu` is recognized within the domain `nps.edu`, but the domain `us.army.mil` does not necessarily know John Doe's address in NPS – nor does `us.army.mil` directly know what his valid roles or attributes are, so the `us.army.mil` file server does not inherently know whether a request coming from John Doe is permissible.

The domain composition problem is encapsulated in *cross-domain access control*. Identity management (*e.g.* account expirations, role assignments, attribute interpretations)

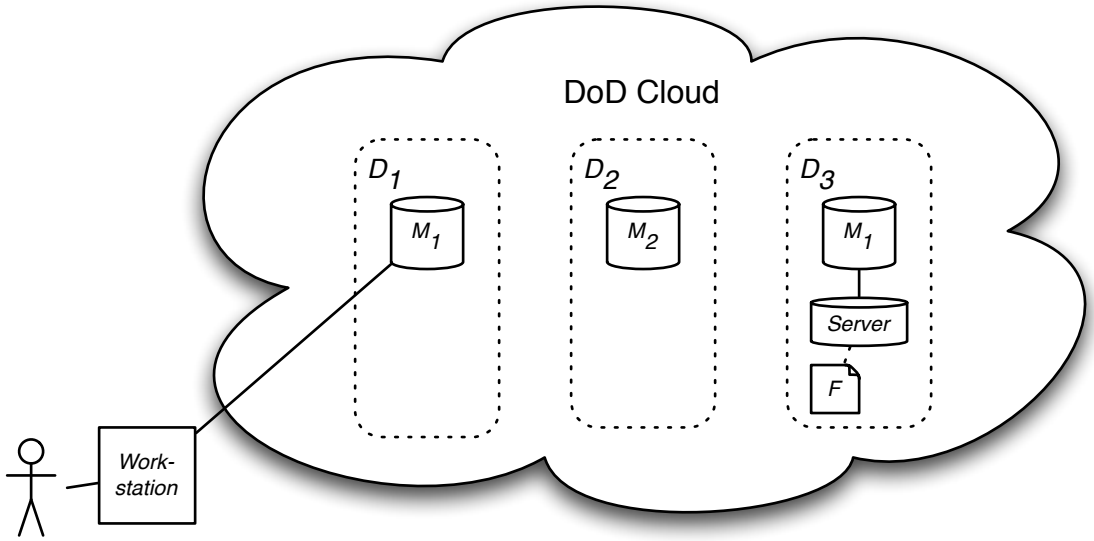

Figure 1: Trust relations among administrative domains. Each domain D_i has a security manager server M_i. A user, registered in domain D_1 but not registered in D_3 has need of file F under D_3, held by *Server*. However, *Server* has no trust relation (pictured with a solid line) with D_1, so any request from D_1 presenting the user's identity is rejected.

occurs at the domain level, but in the Cloud users are likely to employ applications or data lying in other domains. How this is done efficiently and scalably is under current research, and this is just one piece of access control in general – the other three pieces are authentication, authorization, and access decisions. Karp *et al.* propose cross-domain resource requests be made with authorizations, versus the present practice of authentications, and they explore an authorization-based mechanism [10]. While there are many other issues in cross-domain security, we focus this section on the identity representation issue: once an identity is to be presented outside its original domain, the identity should be recognized Cloud-wide.

If there is an *identity-representative trust relation* (referred to as a "trust relation" between organizations) established between two domains, *e.g.* `us.army.mil` and `nps.edu`, then the two domains trust each other to manage and present identities and roles at least as well as to themselves. The domains can negotiate this level of trust and capabilities between each other. However, this approach does not scale if done pairwise between all domains within the Cloud, as noted in Section 1.

Indirect trust relations need to be addressed at a DoD policy level. Figure 2 illustrates one possible policy decision, the *transitive-closure policy*. In the example, the same user in D_1 wishing to access the same file in D_3. The two domains don't have a trust relation with each other, but they do have a trust relation with D_2. With the transitive-closure

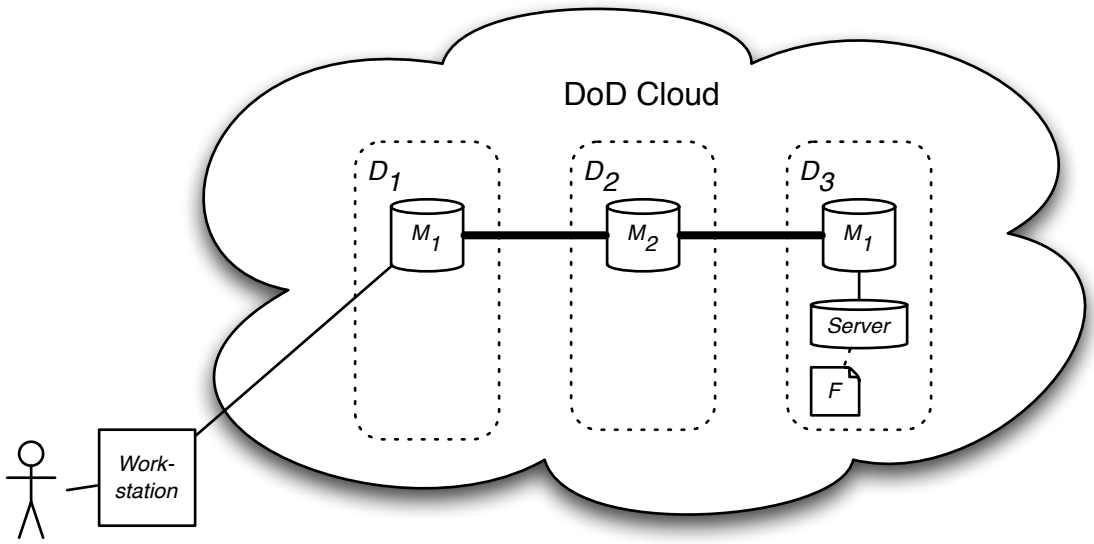

Figure 2: Transitive trust relations among administrative domains. The DoD would need to specify as policy that domains may use intermediary domains to present identity credentials.

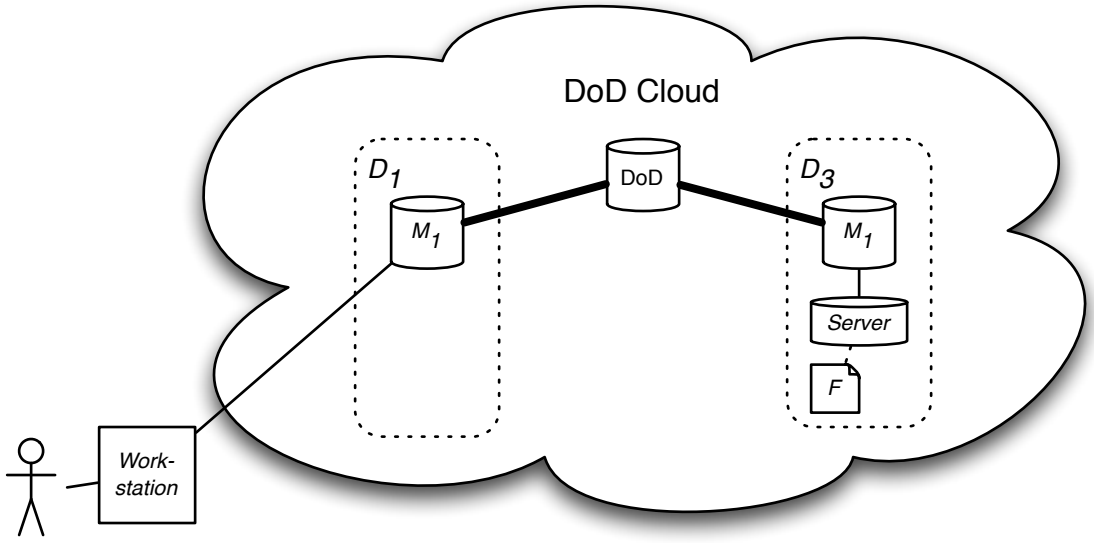

Figure 3: Hierarchical trust relations among administrative domains. The DoD would need to specify that any domain that is a member of the Cloud is capable of presenting identities and roles at least as well as all other domains within the Cloud.

policy, organizations choose whom they trust to present identities, and identities can be forwarded along those chains of trust. This policy would allow organizations to restrict other organizations' access to their data, *by entire organizations*, as policy.

A downside to this policy is that there are other mechanisms to restrict access which are more appropriate than ignoring an identity representation service. This policy also potentially induces convoluted graph traversals as personnel attempt to find whether two organization are reachable in the trust-relation-reachable set. (A social-graph-traversing mechanism for organizations would likely have to become part of the Cloud in this case.) The more frequent organizational policy decisions would be simply whether to establish trust relations and permit transitivity. This approach is likely to not scale to the entire body of organizations in the DoD.

An alternative policy decision takes advantage of extant infrastructure. Figure 3 illustrates the *hierarchical trust relation* policy. In this policy, each organization that joins the Cloud adheres to the same minimum set of standards related to identity services, such as personnel vetting and role assignments. If the DoD policy mandates that certain identity and role representation practices be a requirement to join the Cloud, then all domains within the DoD can be trusted to present identities at least as well as one another.

We believe that for identity recognition, the hierarchical policy is preferable to the transitive-closure policy. For one, identity representation should be within the minimum policy specification from DoD for Cloud usage. For another, much of the identity infrastructure already exists: The PKI infrastructure already used in DoD presents a ready-to-use hierarchy, and the Common Access Card (CAC) infrastructure is already well understood by the DoD workforce. Finally, the transitive-closure policy presents organizational decisions that can quickly become politically charged: The transitive-closure policy partitions DoD into *cliques* of organizations – groups of organizations that are able to agree on identities only with other members of the clique. Adding other trust relations establishes bridges between cliques, which inadvertently joins many organizations together. Joining cliques together then becomes a large negotiating burden on the agencies attempting to establish mutual identity-recognition services.

We propose that for access control purposes, DAC rules are the point where resource access decisions are made and enforced – not identity recognition. It is up to the DoD to decide whether there are other reasons that factor into the identity resolution problem.

2.4 Single sign-on

In the current desktop workflow, users often need to authenticate to many services. Starting after logging on to their desktop and network, they enter a password for email, others for web-based services, and perhaps more for other client-server based applications. The intent of single sign-on is to reduce the number of authenticating actions the user must make to use Cloud resources down to one or two.

A reasonable minimum of authentication actions is two, for the entire time the user is

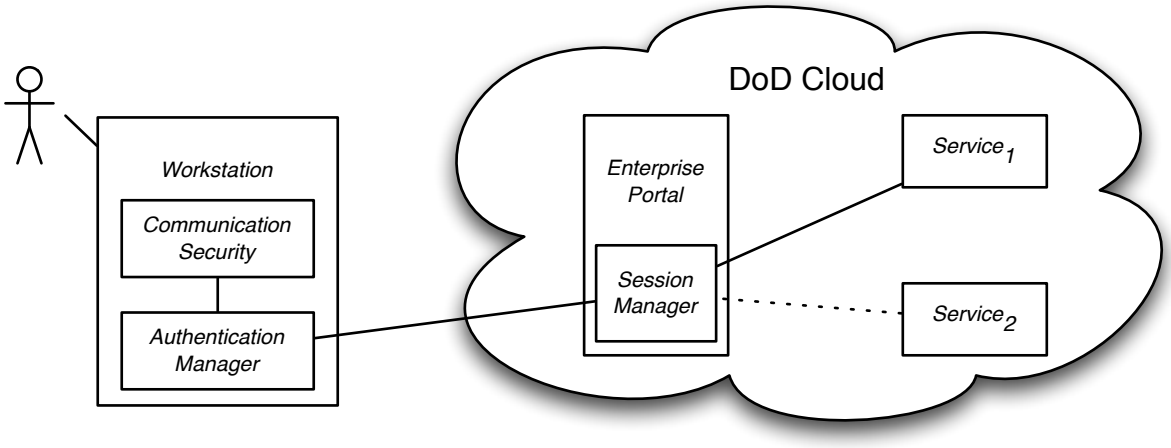

Figure 4: Communication for a user's work session. The user authenticates to a session manager on the Cloud via the local workstation. On successful authentication, the user then possesses a security token representing their work session, that can be forwarded to other Cloud applications. Here, we illustrate a session with lines representing security communications, the solid lines representing active use. $Service_1$ is currently employed and $Service_2$ will soon be a part of this user's session, without requiring an extra authenticating action (*e.g.* entering a CAC PIN) from the user. Note: the location of applications, infrastructure, *etc.* are transparent to the user.

sitting at their desk, though one is the goal. Within the Cloud, the user should only have to authenticate once, which would be single sign-on scoped to the Cloud. Authenticating to the physical workstation or that workstation's network may have to be a separate authentication. No more than those should be necessary for a self-contained system like the Cloud; if there are more, it could be detrimental to productivity[2].

The Cloud should be able to enable single sign-on by treating the user's interaction as in the already-pervasive *session* network protocol. To illustrate, Figure 4 presents an authentication model for a Cloud *work session*. When the user signs on to the Cloud, their work session begins in an Enterprise Portal displaying their available applications and perhaps other "Dashboard-style" status reports. A user's authentication to the Cloud is represented by a security token, established between the user, the workstation and a Cloud session manager. Enduring throughout the user's work session, this token is then forwarded automatically to whichever service the user requests from the Portal's landing page, such as an email application.

[2]Herley discusses the economic costs of interactions with security mechanisms, with this among his conclusions: over the entire U.S. population, the threshold of time spent on security interactions before the U.S. economy is damaged is about one second per user per day, assuming an exploit that affects 1% of users and costs a total of 10 hours to clean up [6]. The amount of interaction to deem "correct" is a debatable subject, but there are some measures for "too much."

Automated session forwarding is not merely an authentication action. A user request to use a particular application is a type of authorization of the application to use or attach to the user's session, providing *end-to-end authorization*. End-to-end authorization has received a lot of attention in the computer systems literature [8]. The implicit authorization here is a significant security concern, a *Service in the Middle* attack, where a malicious service intercepts a user's credentials and attempts to use them for other purposes.

To ensure that the applications sending and receiving tokens are acting appropriately, instead of taking other actions on behalf of an unsuspecting user, a Clark-Wilson [2] style reference monitor must be in place for this authentication forwarding mechanism. This monitor validates transactions not visible to the user and provides some measure of integrity to Cloud operations. The DoD must mandate that a monitor of this sort be in place, so the Cloud can ease authentication yet still prevent hijacked authorizations.

3 Industry efforts

Industry must supply the infrastructure for the Cloud's user administration and storage. The identity infrastructure exists in part, but requires much integration. Industry will have to provide the following:

- Single sign-on within the Cloud system
- An identity management infrastructure
- A large, distributed, versioned file system
- Resiliency to bandwidth spikes, on the scale of all users in a time zone signing on at the beginning of the workday

3.1 Single sign-on and identity management

Single sign-on, described in Section 2.4, entails having an accepted identity infrastructure. The present DoD PKI infrastructure, including the CAC, can serve as the identity infrastructure for the Cloud. The DoD policy-level decision on transitivity *vs.* hierarchy for indirect trust relations (Section 2.3) impacts how the infrastructure is used – for instance, the transitivity policy could be embodied in allowing cliques to select which intermediary certificate authorities to trust.

With the CAC infrastructure already present, DoD and industry must work together to develop a session manager and reference monitor service. This service is a central piece of the Cloud's application security, so it needs to be built early on in the adoption of Cloud Computing.

3.2 The Cloud file system

Building a distributed file system that can support an estimated 3.5 million users of unclassified DoD computing presents significant challenges in infrastructure management. The Cloud storage is to host data and virtual machine environments, so network reliability will be a paramount concern. Storage security will also provide some application usability challenges.

The Cloud's networking must be able to withstand the *train arrival* problem, where an entire time zone's worth of users signs on to the Cloud simultaneously and sets up their workstations. Recent engineering and research efforts, such as applications of the BitTorrent protocol [5] and de-duplication techniques [17], may be able to lessen some of the workload imposed by the correlation of network traffic. Some bandwidth savings have come from prior research on remote file systems which send only updated file portions in network writes, instead of whole files [14]. Version control software operates in a similar manner, updating files with differences since the last version. Many software ideas are available, and with the software design and hardware supply, industry can address these bandwidth spikes.

An early, visible concern for all joining organizations is operating on Cloud data in a disconnected state. For example, if the domain `nps.edu` is physically disconnected from the Internet by accident or malice, the workers at NPS still have to do their best to work with data within `cloud://nps.edu` as it is disconnected from the rest of the Cloud. More commonly, some workers will likely wish to do some work "Offline," syncing their work back to the Cloud when they are next able to get a network connection. Versioned storage can aid this greatly, as the update history is visible on reconnecting. This use of versioning is common in disconnected software development.

An integral part of the file system is an application programming interface (API) for setting security definitions. This API must implement passive, automatic checks that an ACL modification adheres to DoD and domain policy. The interface also must be caller-agnostic (*e.g.* be able to interact with multiple email client applications), to prevent vendor lock-in. Finally, versioning within the file system should provide for both accidental deletion protection[3], and for "Lazy" revocation[4]. Industry should provide an effective interface for versioning, likely drawn from research and production file systems that provide versioning systems, and source code management systems.

4 Milestones for an organization joining the DoD Cloud

We outline here some more concrete steps for individual organizations to take towards this milestone, presenting logistical and technical challenges in a migration to the Cloud as

[3] "A [snapshot] is useful even if it is kept for just a few hours, because users usually notice immediately when they have removed an important file" [7].

[4] "Lazy" revocation revokes access to a file at and after a certain time.

outlined before [4]. The migration we present includes a "Staging" phase for organizations that will join the Cloud.

4.1 Milestone: Virtualizing local storage

This milestone is complete when user desktops only use their local hard drives (if any) as caching devices necessary for local computation. The primary storage is in a data center, perhaps small and local to the building. Getting to this point involves deploying a mandatory security architecture.

The virtual desktop environment is a major, early milestone for transition. Because one goal is more carefully administering the storage infrastructure, an intermediary milestone is acclimating users to using no local storage. Providing, or even mandating, the use of network shared volumes gets users to adjust their workflows to the new storage paradigm.

On the technical-administrative side, we propose that all this storage will *not* start in the Cloud. For an office of 1,000 users, it is reasonable to start by using a storage array local to the building. As the transition to a federal cloud is completed, this storage array can be repurposed as a local cache. Before that transition occurs, this gives the local administrative group its chance to work out reliability and other issues, *e.g.* referencing data within a `cloud://` URI scheme.

4.2 Milestone: Adapting security infrastructure

This milestone is complete when a single sign-on system for desktop use, storage and application access is deployed. The DoD must complete its specification for the minimum amount of personnel information to be centrally reported and stored before this can start, but it will support user administration with strong influence from personnel management local to the organization.

Part of the transition to a Cloud is joining user-administrative information to a global DoD personnel database. It is thus important for the user management information to normalize to the identification and role definition requirements that will be spread throughout DoD.

4.3 Milestone: Virtualize workflow and software

This milestone is complete when users are no longer bound to the applications or operating system state on any particular workstation; work can even be on virtualized machines. It also requires adaption of software presently critical to the organization, from desktop applications (if any) to an interface accessible through the Cloud. This is equivalent to the end of the Migration stage on the path to Cloud Nirvana [4].

Migrating data to the Cloud provides an opportunity for a revolution in improved data integrity and versioning. Some applications will benefit from a well-specified data layout, which can be specified with various software engineering principles [3]. If individual

records are treated as objects, aggregations of those records can benefit from computer science, databases and security research. Updating the records can follow a machine-enforceable pattern, which grants automatic integrity protections, benefiting data security in the manner of the Clark-Wilson security model of data integrity [2].

If aggregations are treated as sets of rules operating on data by reference, this enables a form of file versioning amenable to distributed and offline work, known to file system [7, 15], forensics [17] and software engineering [1] researchers and practitioners. Versioned file system research enables space-efficient definitions of file systems at different points in time by storing base images and small revisions. Distributed source control, such as in the source code manager that enables distributed development of the Linux kernel [1], also takes advantage of space-efficient data versioning (which has a side benefit of network efficiency). Given an appropriately vetted data re-integration application, the Cloud can allow for disconnected work when necessary. Recasting applications with these recommendations is the Integration stage on the path to Cloud Nirvana, and completing and evolving the applications is the final Unification stage [4].

4.4 Milestone: Joining building-local storage to the Cloud

Given the storage infrastructure acquired for a graduated integration, organizations that take this approach will have an important array of storage available for a local cache. Joining the Cloud will provide the last integration steps:

- Storage will have redundancy within the Cloud.
- Cloud applications, such as for administration, will be able to ease operations DoD wide with a unified data store.
- Cloud DAC and reference passing will be fully available.

5 Conclusion

We outlined in this document steps for the three major enablers of the DoD Cloud. The DoD policy specifications provide security interaction direction and high-level specifications for storage and sharing. The industry recommendations guide development efforts in software and hardware infrastructure. Finally, the organization recommendations note the progression towards joining the Cloud and unifying storage, applications and collaborations across the DoD.

6 Acknowledgements

We thank Man-Tak Shing and Loren Peitso for their assistance in improving this paper.

THIS PAGE INTENTIONALLY LEFT BLANK

References

[1] Git - fast version control system. http://git-scm.com/.

[2] David D. Clark and David R. Wilson. A comparison of commercial and military computer security policies. In *IEEE Symposium on Security and Privacy*, pages 184–194, Los Alamitos, CA, USA, 1987. IEEE Computer Society.

[3] Doron Drusinsky, James Bret Michael, Thomas W. Otani, and Man-Tak Shing. Putting order into the cloud: Object-oriented UML-based rule enforcement for document and application organization. Technical Report NPS-CS-10-009, Naval Postgraduate School, Monterey, CA, September 2010.

[4] Kevin D. Foster, John J. Shea, Doron Drusinsky, James Bret Michael, Thomas W. Otani, and Man-Tak Shing. Removing the boundaries: Steps toward a cloud nirvana. In *Proceedings of the 2010 IEEE International Conference on Granular Computing*, pages 167–171, Los Alamitos, CA, USA, August 2010. IEEE Computer Society.

[5] Larry Gadea and Matt Freels. Murder. `http://github.com/lg/murder`, 2010.

[6] Cormac Herley. So long, and no thanks for the externalities: the rational rejection of security advice by users. In *Proceedings of the 2009 Workshop on New Security Paradigms*, NSPW '09, pages 133–144, New York, NY, USA, 2009. ACM.

[7] Dave Hitz, James Lau, and Michael Malcolm. File system design for an NFS file server appliance. In *WTEC'94: Proceedings of the USENIX Winter 1994 Technical Conference*, Berkeley, CA, USA, 1994. USENIX Association.

[8] Jon Howell and David Kotz. End-to-end authorization. In *OSDI'00: Proceedings of the 4th Symposium on Operating System Design & Implementation*, Berkeley, CA, USA, 2000. USENIX Association.

[9] M.L. Johnson, S.M. Bellovin, R.W. Reeder, and S.E. Schechter. Laissez-faire file sharing. In *New Security Paradigms Workshop (NSPW)*, 2009.

[10] Alan H. Karp, Harry Haury, and Michael H. Davis. From ABAC to ZBAC: The evolution of access control models. Technical Report HPL-2009-30, HP Laboratories, 2009.

[11] Butler W. Lampson. Protection. *SIGOPS Operating Systems Review*, 8(1):18–24, 1974.

[12] Andrew W. Leung, Shankar Pasupathy, Garth Goodson, and Ethan L. Miller. Measurement and analysis of large-scale network file system workloads. In *USENIX 2008 Annual Technical Conference*, pages 213–226, Berkeley, CA, USA, 2008. USENIX Association.

[13] Stefan Miltchev, Jonathan M. Smith, Vassilis Prevelakis, Angelos Keromytis, and Sotiris Ioannidis. Decentralized access control in distributed file systems. *ACM Computing Surveys*, 40(3):1–30, 2008.

[14] Athicha Muthitacharoen, Benjie Chen, and David Mazières. A low-bandwidth network file system. In *Proceedings of the Eighteenth ACM Symposium on Operating Systems Principles (SOSP '01)*, pages 174–187, New York, NY, USA, 2001. ACM.

[15] Zachary Peterson and Randal Burns. Ext3cow: a time-shifting file system for regulatory compliance. *Transactions on Storage*, 1(2):190–212, 2005.

[16] Ashlee Vance. WikiLeaks struggles to stay online after attacks. `http://www.nytimes.com/2010/12/04/world/europe/04domain.html`, December 2010.

[17] Kathryn Watkins, Mike McWhorte, Jeff Long, and Bill Hill. Teleporter: An analytically and forensically sound duplicate transfer system. *Digital Investigation*, 6(Supplement 1):S43 – S47, 2009. The Proceedings of the Ninth Annual DFRWS Conference.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Research Sponsored Programs Office, Code 41
   Naval Postgraduate School
   Monterey, CA 93943

4. Professor Peter Denning
   Naval Postgraduate School
   Monterey, California

5. Professor Doron Drusinsky
   Naval Postgraduate School
   Monterey, California

6. Professor Bret Michael
   Naval Postgraduate School
   Monterey, California

7. Professor Thomas Otani
   Naval Postgraduate School
   Monterey, California

8. Professor Man-Tak Shing
   Naval Postgraduate School
   Monterey, California

9. Mr. John Shea
   Office of the DoD Chief Information Officer
   Arlington, Virginia

10. COL Kevin Foster, USA
    Office of the DoD Chief Information Officer
    Arlington, Virginia

11. Professor George Dinolt
    Naval Postgraduate School
    Monterey, California

12. Professor Loren Peitso
    Naval Postgraduate School
    Monterey, California

13. Professor Scott Cote
    Naval Postgraduate School
    Monterey, California

14. Professor Albert Barreto
    Naval Postgraduate School
    Monterey, California

15. Mr. Alex Nelson
    Naval Postgraduate School
    Monterey, California

16. Mr. Scott J Dowell
    Computer Science Corporation
    San Diego, California

17. Mr. Michael Lee
    Touchstone Consulting Group
    Washington, D.C.

18. Dr. Karen Gordon
    Institute for Defense Analyses
    Alexandria, Virginia

19. Dr. Jeffrey Voas
    National Institute of Standards and Technology
    Gaithersburg, Maryland

20. Dr. Mark Lee Badger
    National Institute of Standards and Technology
    Gaithersburg, Maryland

21. Dr. Tim Grance
    National Institute of Standards and Technology
    Gaithersburg, Maryland